

1/11

SEQUENCE LISTING

<110> TANABE SEIYAKU CO., LTD.

<120> Three-dimensional structure of dipeptidyl peptidase IV

<130> 03-039-PCT

<150> US 60/398,761

<151> 2002-07-29

<160> 2

<170> PatentIn version 3.1

<210> 1

<211> 2301

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(2301)

<223>

<400> 1

atg aag aca ccg tgg aag gtt ctt ctg gga ctg ctg ggt gct gct gcg

48

Met	Lys	Thr	Pro	Trp	Lys	Val	Leu	Leu	Gly	Leu	Leu	Gly	Ala	Ala	Ala		
1			5					10					15				
ctt	gtc	acc	atc	atc	acc	gtg	ccc	gtg	ggt	ctg	ctg	aac	aaa	ggc	aca	96	
Leu	Val	Thr	Ile	Ile	Thr	Val	Pro	Val	Val	Leu	Leu	Asn	Lys	Gly	Thr		
			20					25					30				
gat	gat	gct	aca	gct	gac	agt	cgc	aaa	act	tac	act	cta	act	gat	tac	144	
Asp	Asp	Ala	Thr	Ala	Asp	Ser	Arg	Lys	Thr	Tyr	Thr	Leu	Thr	Asp	Tyr		
			35					40					45				
tta	aaa	aat	act	tat	aga	ctg	aag	tta	tac	tcc	tta	aga	tgg	att	tca	192	
Leu	Lys	Asn	Thr	Tyr	Arg	Leu	Lys	Leu	Tyr	Ser	Leu	Arg	Trp	Ile	Ser		
			50					55					60				
gat	cat	gaa	tat	ctc	tac	aaa	caa	gaa	aat	aat	atc	tig	gta	ttc	aat	240	
Asp	His	Glu	Tyr	Leu	Tyr	Lys	Gln	Glu	Asn	Asn	Ile	Leu	Val	Phe	Asn		
65					70					75				80			
gct	gaa	tat	gga	aac	agc	tca	gtt	ttc	tig	gag	aac	agt	aca	ttt	gat	288	
Ala	Glu	Tyr	Gly	Asn	Ser	Ser	Val	Phe	Leu	Glu	Asn	Ser	Thr	Phe	Asp		
					85					90				95			
gag	ttt	gga	cat	tct	atc	aat	gat	tat	tca	ata	tct	ccg	gat	ggg	cag	336	
Glu	Phe	Gly	His	Ser	Ile	Asn	Asp	Tyr	Ser	Ile	Ser	Pro	Asp	Gly	Gln		
			100					105					110				
ttt	att	ctc	tta	gaa	tac	aac	tac	gtg	aag	caa	tgg	agg	cat	tcc	tac	384	
Phe	Ile	Leu	Leu	Glu	Tyr	Asn	Tyr	Val	Lys	Gln	Trp	Arg	His	Ser	Tyr		
			115					120					125				
aca	gct	tca	tat	gac	att	tat	gat	tta	aat	aaa	agg	cag	ctg	att	aca	432	
Thr	Ala	Ser	Tyr	Asp	Ile	Tyr	Asp	Leu	Asn	Lys	Arg	Gln	Leu	Ile	Thr		
			130					135					140				
gaa	gag	agg	att	cca	aac	aac	aca	cag	tgg	gtc	aca	tgg	tca	cca	gtg	480	

[illegible]

4/11

Ala Pro Ala Ser Met Leu Ile Gly Asp His Tyr Leu Cys Asp Val Thr	
290	295
300	
tgg gca aca caa gaa aga att tct ttg cag tgg ctc agg agg att cag	960
Trp Ala Thr Gln Glu Arg Ile Ser Leu Gln Trp Leu Arg Arg Ile Gln	
305	310
315	320
aac tat tgc gtc atg gat att tgt gac tat gat gaa tcc agt gga aga	1008
Asn Tyr Ser Val Met Asp Ile Cys Asp Tyr Asp Glu Ser Ser Gly Arg	
325	330
335	
tgg aac tgc tta gtg gca cgg caa cac att gaa atg agt act act ggc	1056
Trp Asn Cys Leu Val Ala Arg Gln His Ile Glu Met Ser Thr Thr Gly	
340	345
350	
tgg gtt gga aga ttt agg cct tca gaa cct cat ttt acc ctt gat ggt	1104
Trp Val Gly Arg Phe Arg Pro Ser Glu Pro His Phe Thr Leu Asp Gly	
355	360
365	
aat agc ttc tac aag atc atc agc aat gaa gaa ggt tac aga cac att	1152
Asn Ser Phe Tyr Lys Ile Ile Ser Asn Glu Glu Gly Tyr Arg His Ile	
370	375
380	
tgc tat ttc caa ata gat aaa aaa gac tgc aca ttt att aca aaa ggc	1200
Cys Tyr Phe Gln Ile Asp Lys Lys Asp Cys Thr Phe Ile Thr Lys Gly	
385	390
395	400
acc tgg gaa gtc atc ggg ata gaa gct cta acc agt gat tat cta tac	1248
Thr Trp Glu Val Ile Gly Ile Glu Ala Leu Thr Ser Asp Tyr Leu Tyr	
405	410
415	
tac att agt aat gaa tat aaa gga atg cca gga gga agg aat ctt tat	1296
Tyr Ile Ser Asn Glu Tyr Lys Gly Met Pro Gly Gly Arg Asn Leu Tyr	
420	425
430	
aaa atc caa ctt agt gac tat aca aaa gtg aca tgc ctc agt tgt gag	1344

5/11

Lys Ile Gln Leu Ser Asp Tyr Thr Lys Val Thr Cys Leu Ser Cys Glu
 435 440 445
 ctg aat ccg gaa agg tgt cag tac tat tct gtg tca ttc agt aaa gag 1392
 Leu Asn Pro Glu Arg Cys Gln Tyr Tyr Ser Val Ser Phe Ser Lys Glu
 450 455 460
 gcg aag tat tat cag ctg aga tgt tcc ggt cct ggt ctg ccc ctc tat 1440
 Ala Lys Tyr Tyr Gln Leu Arg Cys Ser Gly Pro Gly Leu Pro Leu Tyr
 465 470 475 480
 act cta cac agc agc gtg aat gat aaa ggg ctg aga gtc ctg gaa gac 1488
 Thr Leu His Ser Ser Val Asn Asp Lys Gly Leu Arg Val Leu Glu Asp
 485 490 495
 aat tca gct ttg gat aaa atg ctg cag aat gtc cag atg ccc tcc aaa 1536
 Asn Ser Ala Leu Asp Lys Met Leu Gln Asn Val Gln Met Pro Ser Lys
 500 505 510
 aaa ctg gac ttc att att ttg aat gaa aca aaa ttt tgg tat cag atg 1584
 Lys Leu Asp Phe Ile Ile Leu Asn Glu Thr Lys Phe Trp Tyr Gln Met
 515 520 525
 atc ttg cct cct cat ttt gat aaa tcc aag aaa tat cct cta cta tta 1632
 Ile Leu Pro Pro His Phe Asp Lys Ser Lys Lys Tyr Pro Leu Leu Leu
 530 535 540
 gat gtg tat gca ggc cca tgt agt caa aaa gca gac act gtc ttc aga 1680
 Asp Val Tyr Ala Gly Pro Cys Ser Gln Lys Ala Asp Thr Val Phe Arg
 545 550 555 560
 ctg aac tgg gcc act tac ctt gca agc aca gaa aac att ata gla gct 1728
 Leu Asn Trp Ala Thr Tyr Leu Ala Ser Thr Glu Asn Ile Ile Val Ala
 565 570 575
 agc ttt gat ggc aga gga agt ggt tac caa gga gat aag atc atg cat 1776

6/11

Ser Phe Asp Gly Arg Gly Ser Gly Tyr Gln Gly Asp Lys Ile Met His	
580 585 590	
gca atc aac aga aga ctg gga aca ttt gaa gtt gaa gat caa att gaa	1824
Ala Ile Asn Arg Arg Leu Gly Thr Phe Glu Val Glu Asp Gln Ile Glu	
595 600 605	
gca gcc aga caa ttt tca aaa atg gga ttt gtg gac aac aaa cga att	1872
Ala Ala Arg Gln Phe Ser Lys Met Gly Phe Val Asp Asn Lys Arg Ile	
610 615 620	
gca att tgg ggc tgg tca tat gga ggg tac gta acc tca atg gtc ctg	1920
Ala Ile Trp Gly Trp Ser Tyr Gly Gly Tyr Val Thr Ser Met Val Leu	
625 630 635 640	
gga tcg gga agt ggc gig ttc aag tgt gga ata gcc gtg gcg cct gta	1968
Gly Ser Gly Ser Gly Val Phe Lys Cys Gly Ile Ala Val Ala Pro Val	
645 650 655	
tcc cgg tgg gag tac tat gac tca gtg tac aca gaa cgt tac atg ggt	2016
Ser Arg Trp Glu Tyr Tyr Asp Ser Val Tyr Thr Glu Arg Tyr Met Gly	
660 665 670	
ctc cca act cca gaa gac aac ctt gac cat tac aga aat tca aca gtc	2064
Leu Pro Thr Pro Glu Asp Asn Leu Asp His Tyr Arg Asn Ser Thr Val	
675 680 685	
atg agc aga gct gaa aat ttt aaa caa gtt gag tac ctc ctt att cat	2112
Met Ser Arg Ala Glu Asn Phe Lys Gln Val Glu Tyr Leu Leu Ile His	
690 695 700	
gga aca gca gat gat aac gtt cac ttt cag cag tca gct cag atc tcc	2160
Gly Thr Ala Asp Asp Asn Val His Phe Gln Gln Ser Ala Gln Ile Ser	
705 710 715 720	
aaa gcc ctg gtc gat gtt gga gtg gat ttc cag gca atg tgg tat act	2208

7/11

Lys Ala Leu Val Asp Val Gly Val Asp Phe Gln Ala Met Trp Tyr Thr

725

730

735

gat gaa gac cat gga ata gct agc agc aca gca cac caa cat ata tat 2256

Asp Glu Asp His Gly Ile Ala Ser Ser Thr Ala His Gln His Ile Tyr

740

745

750

acc cac atg agc cac ttc ata aaa caa tgt ttc tct tta cct tag 2301

Thr His Met Ser His Phe Ile Lys Gln Cys Phe Ser Leu Pro

755

760

765

<210> 2

<211> 766

<212> PRT

<213> Homo sapiens

<400> 2

Met Lys Thr Pro Trp Lys Val Leu Leu Gly Leu Leu Gly Ala Ala Ala

1

5

10

15

Leu Val Thr Ile Ile Thr Val Pro Val Val Leu Leu Asn Lys Gly Thr

20

25

30

Asp Asp Ala Thr Ala Asp Ser Arg Lys Thr Tyr Thr Leu Thr Asp Tyr

35

40

45

Leu Lys Asn Thr Tyr Arg Leu Lys Leu Tyr Ser Leu Arg Trp Ile Ser

50

55

60

Asp His Glu Tyr Leu Tyr Lys Gln Glu Asn Asn Ile Leu Val Phe Asn

65

70

75

80

8/11

Ala Glu Tyr Gly Asn Ser Ser Val Phe Leu Glu Asn Ser Thr Phe Asp
85 90 95

Glu Phe Gly His Ser Ile Asn Asp Tyr Ser Ile Ser Pro Asp Gly Gln
100 105 110

Phe Ile Leu Leu Glu Tyr Asn Tyr Val Lys Gln Trp Arg His Ser Tyr
115 120 125

Thr Ala Ser Tyr Asp Ile Tyr Asp Leu Asn Lys Arg Gln Leu Ile Thr
130 135 140

Glu Glu Arg Ile Pro Asn Asn Thr Gln Trp Val Thr Trp Ser Pro Val
145 150 155 160

Gly His Lys Leu Ala Tyr Val Trp Asn Asn Asp Ile Tyr Val Lys Ile
165 170 175

Glu Pro Asn Leu Pro Ser Tyr Arg Ile Thr Trp Thr Gly Lys Glu Asp
180 185 190

Ile Ile Tyr Asn Gly Ile Thr Asp Trp Val Tyr Glu Glu Glu Val Phe
195 200 205

Ser Ala Tyr Ser Ala Leu Trp Trp Ser Pro Asn Gly Thr Phe Leu Ala
210 215 220

Tyr Ala Gln Phe Asn Asp Thr Glu Val Pro Leu Ile Glu Tyr Ser Phe
225 230 235 240

Tyr Ser Asp Glu Ser Leu Gln Tyr Pro Lys Thr Val Arg Val Pro Tyr
245 250 255

Pro Lys Ala Gly Ala Val Asn Pro Thr Val Lys Phe Phe Val Val Asn
260 265 270

Thr Asp Ser Leu Ser Ser Val Thr Asn Ala Thr Ser Ile Gln Ile Thr
275 280 285

Ala Pro Ala Ser Met Leu Ile Gly Asp His Tyr Leu Cys Asp Val Thr

9/11

290	295	300	
Trp Ala Thr Gln Glu Arg Ile Ser Leu Gln Trp Leu Arg Arg Ile Gln			
305	310	315	320
Asn Tyr Ser Val Met Asp Ile Cys Asp Tyr Asp Glu Ser Ser Gly Arg			
	325	330	335
Trp Asn Cys Leu Val Ala Arg Gln His Ile Glu Met Ser Thr Thr Gly			
	340	345	350
Trp Val Gly Arg Phe Arg Pro Ser Glu Pro His Phe Thr Leu Asp Gly			
	355	360	365
Asn Ser Phe Tyr Lys Ile Ile Ser Asn Glu Glu Gly Tyr Arg His Ile			
	370	375	380
Cys Tyr Phe Gln Ile Asp Lys Lys Asp Cys Thr Phe Ile Thr Lys Gly			
385	390	395	400
Thr Trp Glu Val Ile Gly Ile Glu Ala Leu Thr Ser Asp Tyr Leu Tyr			
	405	410	415
Tyr Ile Ser Asn Glu Tyr Lys Gly Met Pro Gly Gly Arg Asn Leu Tyr			
	420	425	430
Lys Ile Gln Leu Ser Asp Tyr Thr Lys Val Thr Cys Leu Ser Cys Glu			
	435	440	445
Leu Asn Pro Glu Arg Cys Gln Tyr Tyr Ser Val Ser Phe Ser Lys Glu			
	450	455	460
Ala Lys Tyr Tyr Gln Leu Arg Cys Ser Gly Pro Gly Leu Pro Leu Tyr			
465	470	475	480
Thr Leu His Ser Ser Val Asn Asp Lys Gly Leu Arg Val Leu Glu Asp			
	485	490	495
Asn Ser Ala Leu Asp Lys Met Leu Gln Asn Val Gln Met Pro Ser Lys			
	500	505	510

10/11

Lys Leu Asp Phe Ile Ile Leu Asn Glu Thr Lys Phe Trp Tyr Gln Met
515 520 525

Ile Leu Pro Pro His Phe Asp Lys Ser Lys Lys Tyr Pro Leu Leu Leu
530 535 540

Asp Val Tyr Ala Gly Pro Cys Ser Gln Lys Ala Asp Thr Val Phe Arg
545 550 555 560

Leu Asn Trp Ala Thr Tyr Leu Ala Ser Thr Glu Asn Ile Ile Val Ala
565 570 575

Ser Phe Asp Gly Arg Gly Ser Gly Tyr Gln Gly Asp Lys Ile Met His
580 585 590

Ala Ile Asn Arg Arg Leu Gly Thr Phe Glu Val Glu Asp Gln Ile Glu
595 600 605

Ala Ala Arg Gln Phe Ser Lys Met Gly Phe Val Asp Asn Lys Arg Ile
610 615 620

Ala Ile Trp Gly Trp Ser Tyr Gly Gly Tyr Val Thr Ser Met Val Leu
625 630 635 640

Gly Ser Gly Ser Gly Val Phe Lys Cys Gly Ile Ala Val Ala Pro Val
645 650 655

Ser Arg Trp Glu Tyr Tyr Asp Ser Val Tyr Thr Glu Arg Tyr Met Gly
660 665 670

Leu Pro Thr Pro Glu Asp Asn Leu Asp His Tyr Arg Asn Ser Thr Val
675 680 685

Met Ser Arg Ala Glu Asn Phe Lys Gln Val Glu Tyr Leu Leu Ile His
690 695 700

Gly Thr Ala Asp Asp Asn Val His Phe Gln Gln Ser Ala Gln Ile Ser
705 710 715 720

Lys Ala Leu Val Asp Val Gly Val Asp Phe Gln Ala Met Trp Tyr Thr

11/11

725	730	735
Asp Glu Asp His Gly Ile Ala Ser Ser Thr Ala His Gln His Ile Tyr		
740	745	750
Thr His Met Ser His Phe Ile Lys Gln Cys Phe Ser Leu Pro		
755	760	765